Upper bounds for the error in some interpolation and extrapolation designs

Michel Broniatowski[1],Giorgio Celant[2], Marco Di Battista[2] , Samuela Leoni-Aubin[3]

Upper bounds for the error in some interpolation and extrapolation designs Michel Broniatowski[1],Giorgio Celant[2], Marco Di Battista[2] , Samuela Leoni-Aubin[3]

[1]Université Pierre et Marie Curie, LSTA,e-mail:michel.broniatowski@upmc.fr

[2]University of Padua, Department of Statistical Sciences,

[3]INSA Lyon, ICJ, e-mail: samuela.leoni@insa-lyon.fr

Abstract

This paper deals with probabilistic upper bounds for the error in functional estimation defined on some interpolation and extrapolation designs, when the function to estimate is supposed to be analytic. The error pertaining to the estimate may depend on various factors: the frequency of observations on the knots, the position and number of the knots, and also on the error committed when approximating the function through its Taylor expansion. When the number of observations is fixed, then all these parameters are determined by the choice of the design and by the choice estimator of the unknown function.

AMS (2010) Classification: 62K05, 41A5

1 Introduction

Consider a function φ defined on some open set $D \subset \mathbb{R}$ and which can be observed on a compact subset S included in D. The problem that we consider is the estimation of this function through some interpolation or extrapolation techniques. This turns out to define a finite set of points s_i in a domain $\widetilde{S}$ included in S and the number of measurement of the function φ at each of these points, that is to define a design $\mathcal{P} := \left\{ (s_i, n_i) \in S \times \mathbb{N},\ i = 0, ..., l,\ \widetilde{S} \subsetneq S \right\}$. The points s_i are called

the *knots*, n_i is the frequency of observations at knot s_i and $l+1$ is the number of knots. The choice of the design $\mathcal{P}$ is based on some optimality criterion. For example, we could choose an observation scheme that minimize the variance of the estimator of φ.

The choice of $\mathcal{P}$ has been investigated by many authors. Hoel and Levine and Hoel ([8] and [9]) considered the case of the extrapolation of a polynomial function with known degree in one and two variables. Spruill, in a number of papers (see [12], [13], [14] and [15]) proposed a technique for the (interpolation and extrapolation) estimation of a function and its derivatives, when the function is supposed to belong to a Sobolev space, Celant (in [4] and [5]) considered the extrapolation of quasi-analytic functions and Broniatowski-Celant in [3] studied optimal designs for analytic functions through some control of the bias.

The main defect of any interpolation and extrapolation scheme is its extreme sensitivity to the uncertainties pertaining to the values of φ on the knots. The largest the number $l+1$ of knots, the more unstable is the estimate. In fact, even when the function φ is accurately estimated on the knots, the estimates of φ or of one of its derivatives $\varphi^{(j)}$ at some point in D may be quite unsatisfactory, due either to a wrong choice of the number of knots or to their location. The only case when the error committed while estimating the values $\varphi(s_i)$ is not amplified in the interpolation procedure is the linear case. Therefore, for any more envolved case the choice of l and (s_i, n_i) must be handled carefully, which explains the wide literature devoted to this subject. For example, if we estimate $\varphi(v)$, $v \in S \diagdown \widetilde{S}$, by $\widehat{\varphi(s_k)} := \varphi(s_k) + \varepsilon(k)$, where $\varepsilon(k)$ denotes the estimation error and $\widetilde{S}$ a Tchebycheff set of points S, we obtain

$$\left|\varphi(v) - \widehat{\varphi(s_k)}\right| \leq \left(\max_k |\varepsilon(k)|\right) \Lambda_l(v, s_k, 0),$$

where $\Lambda_l(v, s_i, j)$ is a function that depends on $\widetilde{S}$, the number of knots and on the order of the derivative that we aim to estimate (here 0), and (see [2] and [10])

$$\max_{k=0,\ldots,l} \Lambda_l(v, s_k, 0) := \frac{1}{l+1} \sum_{k=0}^{l} ctg\left(\frac{2k+1}{4(l+1)}\pi\right) \sim \frac{2}{\pi} \ln(l+1) \qquad \text{when } l \to \infty.$$

If equidistant knots are used, one gets (see [11])

$$\max_{k=0,\ldots,l} \Lambda_l(v, s_k, 0) \sim \frac{2^{l+1}}{el(\ln l + \gamma)}, \qquad \gamma = 0,577 \text{ (Euler-Mascheroni constant)}.$$

When the bias in the interpolation is zero, as in the case when φ is polynomial with known degree, the design is optimized with respect to

the variance of the interpolated value (see [8]). In the other cases the criterion that is employed is the minimal MSE criterion. The minimal MSE criterion allows the estimator to be as accurate as possible but it does not yield any information on the interpolation/extrapolation error.

In this paper, we propose a probabilistic tool (based on the concentration of measure) in order to control the estimation error. In Section 2 we present the model, the design and the estimators. Section 3 deals with upper bouns for the error. Concluding remarks are given in Section 4.

2 The model, the design and the estimators

Consider an unknown real-valued analytic function f defined on some interval D :

$$\begin{aligned} f: \quad D := (a,b) &\to \mathbb{R} \\ v &\mapsto f(v). \end{aligned}$$

We assume that this function is observable on a compact subset S included in D, $S := [\underline{s}, \overline{s}] \subset D$, and that its derivatives are not observable at any point of D. Let $\widetilde{S} := \left\{ s_k \in \widetilde{S}, k = 0, ..., l \right\}$ be a finite subset of $l+1$ elements in the set S. The points s_k are called the *knots*.
Observations Y_i, $i = 1, \ldots, n$ are generated from the following location-scale model

$$\begin{aligned} Y_j(s_k) =& f(s_k) + \sigma E(Z_j) + \varepsilon_j, \\ \varepsilon_j :=& \sigma Z_j - \sigma E(Z_j), \qquad j = 1, ..., n_k, \quad k = 0, ..., l, \end{aligned}$$

where Z is a completely specified continuous random variable, the location parameter $f(v)$ and the scale parameter $\sigma > 0$ are unknown parameters. $E(Z), \varsigma$ respectively denote the mean and the variance of Z, and n_k is the *frequency of observations* at knot s_k.
We assume to observe $(l+1)$ i.i.d. samples, $\underline{Y}(k) := (Y_1(n_k), ..., Y_{n_k}(n_k)), k = 0, ..., l$, and $Y_i(n_k)$ i.i.d. $Y_1(n_k)$, for all $i \neq k$, $i = 0, \ldots, l$.
The aim is to estimate a derivative of $f(v)$, $f^{(d)}(v)$, $d \in \mathbb{N}$, at a point $v \in (a, \overline{s})$.
Let $\varphi(v) := f(v) + \sigma E(Z)$, and consider the Lagrange polynomial

$$L_{s_k}(v) := \prod_{j \neq k, j=0}^{l} \frac{v - s_j}{s_k - s_j}.$$

We are interested in interpolating (or extrapolating) some derivatives of

φ, $\varphi^{(d)}$, with $d \in \mathbb{N}$,

$$\mathcal{L}_l\left(\varphi^{(d)}\right)(v) := \sum_{k=0}^{l} \varphi(s_k)\, L_{s_k}^{(d)}(v)\,.$$

The domain of extrapolation is denoted $U := D \diagdown S$. It is convenient to define a generic point $v \in D$ stating that it is an *observed point* if it is a knot, an *interpolation point* if $v \in S$ and an *extrapolation point* if $v \in U$. For all $d \in \mathbb{N}$, for any $v \in S$, the Lagrange interpolation scheme converges for the function $\varphi^{(d)}$, that is, for $l \to \infty$,

$$\mathcal{L}_l\left(\varphi^{(d)}\right)(s) \to \varphi^{(d)}(s)\,, \quad \forall s \in S.$$

Interpolating the derivative $\varphi^{(d+i)}(s^*)$ at a point $s^* \in S$ opportunely chosen, a Taylor expansion with order $(m-1)$ of $\varphi^{(d)}(v)$ at point v from s^* gives

$$T_{\varphi^{(d)},m,l}(v) := \sum_{i=0}^{m-1} \frac{(v-s^*)^i}{i!} \mathcal{L}_l\left(\varphi^{(d+i)}\right)(s^*)\,, \quad s^* \in S,$$

and we have

$$\lim_{m\to\infty} \lim_{l\to\infty} T_{\varphi^{(d)},m,l}(v) = \varphi^{(d)}(v)\,, \quad \forall v \in D.$$

When $\varphi^{(d)} \in \mathcal{C}^{\alpha}(D)$, $\forall \alpha$, $l \geq 2\alpha - 3$, the upper bound for the error of approximation is given in [1],

$$E_t := \sup_{v\in D} \left|\varphi^{(d)}(v) - T_{\varphi^{(d)},m,l}(v)\right| \leq M(m,l,\alpha),$$

where $M(m,l,\alpha) = A(\alpha,l) + B(m)$,

$$A(\alpha,l) := K(\alpha,l) \sum_{i=0}^{m-1} \left(\sup_{s\in S} \left|\varphi^{(d+i+\alpha)}(s)\right| \frac{1}{i!} \sup_{v\in U} |v-s^*|^i \right),$$

$$K(\alpha,l) := \left(\frac{\pi}{2(1+l)} (\overline{s} - \underline{s}) \right)^{\alpha} \left(9 + \frac{4}{\pi} \ln(1+l) \right),$$

$$\text{and} \qquad B(m) := \sup_{v\in(a,\overline{s})} \left(\frac{|u-s^*|^m \left|\varphi^{(d+\alpha)}(v)\right|}{m!} \right).$$

The optimal design writes $\left\{ (n_k, s_k) \in (\mathbb{N} \setminus \{0\})^{l+1} \times \mathbb{R}^{l+1},\ n := \sum_{k=0}^{l} n_k,\ n \text{ fixed} \right\}$, where n is the total number of experiments and the $(l+1)$ knots are defined by

$$s_k := \frac{\overline{s} + \underline{s}}{2} - \frac{\overline{s} - \underline{s}}{2} \cos \frac{2k-1}{2l+2}\pi, \qquad k = 0, \ldots, l,$$

with $n_k := \left[\frac{n\sqrt{P_k}}{\sum_{k=0}^{l}\sqrt{P_k}}\right]$, $[.]$ denoting the integer part function, and (see [3] for details)

$$P_k := \left|\sum_{\beta=0}^{m}\sum_{\alpha=0}^{m}\frac{(u-s)^{\alpha+\beta}}{\alpha!\beta!}L_{s_k}^{(\alpha)}(s)\,L_{s_k}^{(\beta)}(s)\right|, \qquad k=0,...,l.$$

The function φ cannot be observed exactly at the knots. Let $\widehat{\varphi(s_k)}$ denote the least squares estimate of $\varphi(s_k)$ at the knot s_k and

$$\mathcal{L}_l\left(\widehat{\varphi^{(d+i)}}\right)(v) := \sum_{k=0}^{l}\widehat{\varphi(s_k)}L_{s_k}^{(d+i)}(v). \tag{1}$$

We estimate the $d-$th derivative of $\varphi(v)$ at $v \in D$ as follows

$$\widehat{T}_{\varphi^{(d)},m,l}(v) := \sum_{i=0}^{m-1}\frac{(v-s^*)^i}{i!}\mathcal{L}_l\left(\widehat{\varphi^{(d+i)}}\right)(s^*), \qquad s^* \in S.$$

The knots s_k are chosen in order to minimize the variance of $\widehat{T}_{\varphi^{(d)},m,l}(v)$ and it holds

$$\lim_{m\to\infty}\lim_{l\to\infty}\lim\nolimits_{\min_{k=0,...,l}(n_k)\to\infty}\widehat{T}_{\varphi^{(d)},m,l}(v) = \varphi^{(d)}(v), \qquad \forall v \in D.$$

$\widehat{T}_{\varphi^{(d)},m,l}(v)$ is an extrapolation estimator when $v \in U$ and an interpolation estimator when $v \in S$.
For a fixed degree l of the Lagrange scheme (1), the total error committed while substituting $\varphi^{(d)}(v)$ by $\widehat{T}_{\varphi^{(d)},m,l}(v)$ writes

$$E_{Tot}\left(\varphi^{(d)}(v)\right) := \varphi^{(d)}(v) - \widehat{T}_{\varphi^{(d)},m,l}(v).$$

For the interpolation error concerning $\varphi^{(i+d)}$, we have the following result presented in [6], p.293 : if $\varphi^{(i+d)} \in \mathcal{C}^{\alpha}(S)$, $\forall\alpha$, $l \geq 2\alpha-3$, then

$$\sup_{s\in S}\left|\varphi^{(d+i)}(s) - \mathcal{L}_l\left(\varphi^{(d+i)}\right)(s)\right| \leq M_1 := K(\alpha,l)\sup_{s\in S}\left|\varphi^{(d+i+\alpha)}(s)\right|.$$

This error depends on the very choice of the knots and is controlled through a tuning of l.
The error due to the Taylor expansion of order $(m-1)$

$$\varphi^{(d)}(v) - \sum_{i=0}^{m-1}\frac{(v-s^*)^i}{i!}\varphi^{(d+i)}(s^*)$$

depends on s^*, it is a truncation error and it can be controlled through a tuning of m.

Let $\widehat{\varphi(s_k)}$ be an estimate of $\varphi(s_k)$ on the knot s_k and

$$\varepsilon\left(k\right) := \varphi\left(s_k\right) - \widehat{\varphi\left(s_k\right)}, \qquad k = 0, ..., l$$

denote the error pertaining to $\varphi(s_k)$ due to this estimation. $\varepsilon\left(k\right)$ clearly depends on n_k, the frequency of observations at knot s_k.
Finally, when n is fixed, the error committed while extrapolating depends on the design $\{(n_k, s_k) \in (\mathbb{N} \setminus \{0\})^{l+1} \times \mathbb{R}^{l+1},\ k = 0, \ldots, l,\ n = \sum_{k=0}^{l} n_k\}$, on m and on l.
Without loss of generality, we will assume $\sigma = 1$. In this case we have $\widehat{\varphi\left(s_k\right)} = \overline{Y}\left(s_k\right) := \frac{\sum_{j=1}^{n_k} Y_j(k)}{n_k}$. The general case when σ is unknown is described in [3].
In the next Section we will provide upper bounds for the errors in order to control them.
Since φ is supposed to be an analytic function, we can consider the extrapolation as an analytic continuation of the function out of the set S obtained by a Taylor expansion from an opportunely chosen point s^* in S. So, the extrapolation error will depend on the order of the Taylor expansion and on the precision in the knowledge of the derivatives of the function at s^*. This precision is given by the interpolation error and by the estimation errors on the knots. The analyticity assumption also implies that the interpolation error will quickly converge to zero. Indeed, for all integer r, the following result holds:

$$\lim_{l \to \infty} l^r \sup_{s \in S} \left| \varphi^{(j)}\left(s\right) - \sum_{k=0}^{l} L_{s_k}^{(j)}\left(s\right) \varphi\left(s_k\right) \right| = 0.$$

We remark that the instability of the interpolation and extrapolation schemes discussed by Runge (1901) can be avoided if the chosen knots form a Tchebycheff set of points in S, or if they form a Feteke set of points in S, or by using splines.
Note that in all the works previously quoted the function is supposed to be polynomial with known degree (in [8] and [9]), to belongs to a Sobolev space (see [12], [13], [14] and [15]), or to be quasi analytic (in [4] and [5]), or analytic (in [3]). Moreover, $\widetilde{S}$ is chosen as a Tchebycheff set of points in S .
Bernstein in [2] affirmed that polynomials of low degree are good approximations for analytic functions. In the case of the Broniatowski-Celant design ([3]), the double approximation to approach φ allows to choose any subset of S as possible interpolation set. So, if the unknown function is supposed to be analytic, then we can choose a small interpolation set in order to obtain a small interpolation error.

3 Upper bounds and control of the error

The extrapolation error depends on three kinds of errors: truncation error, interpolation error and error of estimation of the function on the knots. In order to control the extrapolation error, we split an upper bound for it in a sum of three terms, each term depending only on one of the three kinds of errors.
In the sequel, we will distinguish two cases: in the first case, we suppose that the observed random variable Y is bounded, in the second case Y is supposed to be a random variable with unbounded support. We suppose that the support is known.

3.1 Case 1: Y is a bounded random variable

If τ_1, τ_2 (assumed known) are such that $\Pr(\tau_1 \leq Y \leq \tau_2) = 1$, it holds $|\varphi(v)| \leq R$, where $R := \max\{|\tau_1|, |\tau_2|\}$. Indeed, $E(Y) = \varphi \in [-R, R]$. Let

$$\varepsilon(k) := \frac{\sum_{j=1}^{n_k} Y_j(k)}{n_k} - \varphi(s_k).$$

The variables $Y_j(k), \forall j = 1, ..., n_k, \forall k = 0, .., l$, are i.i.d., with the same bounded support and for all $k, E(Y_j(k)) = \varphi(s_k)$, hence we can apply the Hoeffding's inequality (in [7]):

$$\Pr\{|\varepsilon(k)| \geq \rho\} \leq 2\exp\left(-\frac{2\rho^2 n_k}{(\tau_2 - \tau_1)^2}\right).$$

In Proposition 1, we give an upper bound for the extrapolation error denoted by E_{ext}. This bound is the sum of the three terms, M_{Taylor}, controlling the error associated to the truncation of the Taylor expansion which defines $\varphi^{(d)}$, M_{interp}, controlling the interpolation error and M_{est}, describing the estimation error on the knots.

Proposition 1 *For all $\alpha \in \mathbb{N} \setminus \{0\}$, if $\varphi^{(i+d)} \in \mathcal{C}^{\alpha}(a, b)$, $l \geq 2\alpha - 3$, then, $\forall u \in U$, $|E_{ext}(u)| \leq M_{Taylor} + M_{interp} + M_{est}$, where*

$$M_{Taylor} := R\frac{(d+m)!}{m!}\left(\frac{s^* - u}{b - a}\right)^m \frac{1}{(b-a)^d},$$

$$K(l, \alpha) := \left(9 + \frac{4}{\pi}\ln(1+l)\right)\left(\frac{\pi}{2(1+l)}\right)^{\alpha},$$

$$M_{interp} := K(l, \alpha)\frac{R}{(\overline{s} - \underline{s})^{d+\alpha}}\sum_{i=0}^{m-1}\left(\frac{s^* - u}{\overline{s} - \underline{s}}\right)^i \frac{(d+i+\alpha)!}{i!},$$

$$\Lambda(l,m) := \sum_{i=0}^{m-1}\sum_{k=0}^{l}\frac{(s^*-u)^i}{i!}\left|L_{s_k}^{(d+i)}(s^*)\right|,$$

$$M_{est} := \Lambda(l,m)\left(\max_{k=0,\dots,l}|\varepsilon(k)|\right).$$

Proof. By using the Cauchy's Theorem on the derivatives of the analytic functions, we obtain

$$\left|\varphi^{(d)}(u)-\widehat{\varphi^{(d)}(u)}\right| = \left|\varphi^{(d)}(u)+\sum_{i=0}^{m-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i-\sum_{i=0}^{m-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i-\widehat{\varphi^{(d)}(u)}\right|$$

$$\leq\left|\varphi^{(d)}(u)-\sum_{i=0}^{m-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i\right|+\left|\sum_{i=0}^{m-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i-\widehat{\varphi^{(d)}(u)}\right|$$

$$\leq\frac{\sup_{v\in U}\left|\varphi^{(d+m)}(v)\right|}{m!}(s^*-u)^m+\left|\sum_{i=0}^{m-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i-\sum_{i=0}^{m-1}\frac{\widehat{\varphi^{(d+i)}(s^*)}}{i!}(u-s^*)^i\right|$$

$$\leq\frac{R(m+d)!}{(b-a)^d m!}\left(\frac{s^*-u}{b-a}\right)^m+\left|\sum_{i=0}^{m-1}\frac{(s^*-u)^i}{i!}\left(\varphi^{(d+i)}(s^*)-\widehat{\varphi^{(d+i)}(s^*)}\right)\right|$$

$$\leq\frac{R(m+d)!}{(b-a)^d m!}\left(\frac{s^*-u}{b-a}\right)^m+\sum_{i=0}^{m-1}\frac{(s^*-u)^i}{i!}\left|\varphi^{(d+i)}(s^*)-\widehat{\varphi^{(d+i)}(s^*)}\right|$$

$$\leq M_{Taylor}+\sum_{i=0}^{m-1}\frac{(s^*-u)^i}{i!}\left|\begin{array}{l}\varphi^{(d+i)}(s^*)-\sum_{k=0}^{l}L_{s_k}^{(d+i)}(s^*)\varphi(s_k)\\+\sum_{k=0}^{l}L_{s_k}^{(d+i)}(s^*)\varphi(s_k)-\widehat{\varphi^{(d+i)}(s^*)}\end{array}\right|$$

$$\leq M_{Taylor}+\sum_{i=0}^{m-1}\frac{(s^*-u)^i}{i!}\left|\varphi^{(d+i)}(s^*)-\sum_{k=0}^{l}L_{s_k}^{(d+i)}(s^*)\varphi(s_k)\right|+$$
$$\sum_{i=0}^{m-1}\sum_{k=0}^{l}\frac{(s^*-u)^i}{i!}L_{s_k}^{(d+i)}(s^*)\left|\varphi(s_k)-\overline{Y}(k)\right|$$

$$\leq M_{Taylor}+\sum_{i=0}^{m-1}\frac{(s^*-u)^i}{i!}K(l,\alpha)\left(\sup_{s\in S}\left|\varphi^{(d+i+\alpha)}(s)\right|\right)$$
$$+\sum_{i=0}^{m-1}\sum_{k=0}^{l}\frac{(s^*-u)^i}{i!}L_{s_k}^{(d+i)}(s^*)\left|\varphi(s_k)-\overline{Y}(k)\right|$$

$$\leq M_{Taylor} + \frac{R}{(\overline{s}-\underline{s})^{d+\alpha}} \sum_{i=0}^{m-1} \frac{(s^*-u)^i}{i!} K(l,\alpha) \frac{(d+i+\alpha)!}{(\overline{s}-\underline{s})^i}$$

$$+ \sum_{i=0}^{m-1} \sum_{k=0}^{l} \frac{(s^*-u)^i}{i!} L_{s_k}^{(d+i)}(s^*) \left|\varphi(s_k) - \overline{Y}(k)\right|$$

$$\leq M_{Taylor} + M_{interp} + \sum_{i=0}^{m-1} \sum_{k=0}^{l} \frac{(s^*-u)^i}{i!} L_{s_k}^{(d+i)}(s^*) \left|\varphi(s_k) - \overline{Y}(k)\right|$$

$$\leq M_{Taylor} + M_{interp} + \left(\max_{k=0,\dots,l} |\varepsilon(k)|\right) \sum_{i=0}^{m-1} \sum_{k=0}^{l} \frac{(s^*-u)^i}{i!} \left|L_{s_k}^{(d+i)}(s^*)\right|$$

$$= M_{Taylor} + M_{interp} + M_{est}.$$

■

Proposition 2 yields the smallest integer such that the error of estimation is not greater than a chosen threshold with a fixed probability.

Proposition 2 $\forall \eta \in [0,1], \forall \rho \in \mathbb{R}^+, \exists n \in \mathbb{N}$ *such that*

$$Pr\left(\max_{k=0,\dots,l} |\varepsilon(k)| \geq \frac{\rho}{\Lambda(l,m)}\right) \leq \eta.$$

Proof. If, $\forall k\ |\varepsilon(k)| \geq \frac{\rho}{\Lambda(l,m)}$, then $\max_{k=0,\dots,l} |\varepsilon(k)| \geq \frac{\rho}{\Lambda(l,m)}$. We have

$$\Pr\left(\max_{k=0,\dots,l} |\varepsilon(k)| \geq \frac{\rho}{\Lambda(l,m)}\right) \leq \prod_{k=0}^{l} \Pr\left(|\varepsilon(k)| \geq \frac{\rho}{\Lambda(l,m)}\right) \leq \prod_{k=0}^{l} 2\exp\left(-\frac{2\rho^2}{(\Lambda(l,m))^2} n_k\right).$$

So, we can choose

$$n^* = \left[\frac{(l+1)\ln 2 - \ln \eta}{2} \left(\frac{\Lambda(l,m)(\tau_2-\tau_1)}{\rho}\right)^2\right].$$

■

Proposition 3 gives an upper bound for the extrapolation error that depends on (l,m,n). We recall that the number of knots $l+1$ controls the interpolation error, m denotes the number of terms used in the Taylor expansion for $\varphi^{(d)}$ and n is the total number of observations used to estimate $\varphi(s_k), k = 0,..,l$. Hence n controls the total estimation error.

Proposition 3 *With the same hypotheses and notations, we have that*

$$\forall (\rho_m, \rho_l, \rho_n) \in \mathbb{R}(\mathbb{R}^+)^3, \quad |E_{ext}(u)| \leq \rho_m + \rho_l + \rho_n$$

with probability η. η *depends on the choice of* (ρ_m, ρ_l, ρ_n), *which depends on* (m,l,n).

Proof. When (ρ_m, ρ_l) is fixed , we can choose (m, l) as the solution of the system:

$$(M_{Taylor}, M_{interp}) = (\rho_m, \rho_l).$$

We end the proof by taking $\rho_n = \frac{\rho}{\Lambda(l,m)}$ and $n = n^*$.

■

In the case of the estimation of $\varphi(u)$ (i.e., when $d = 0$) we obtain for the couple (m, n) the explicit solution

$$m = \frac{\ln \rho_m - \ln R}{\ln(s^* - u) - \ln(b - a)},$$

$$n = \left[\frac{(l+1)\ln 2 - \ln \eta}{2} \left(\frac{\Lambda(l)(\tau_2 - \tau_1)}{\rho} \right)^2 \right], \Lambda(l) = \sum_{i=0}^{m-1} \sum_{k=0}^{l} \frac{(s^* - u)^i}{i!} \left| L_{s_k}^{(i)}(s^*) \right|.$$

When $l \geq 2\alpha - 3$, l is the solution of the equation

$$\rho_l = \left(9 + \frac{4}{\pi} \ln(1+l) \right) \left(\frac{\pi}{2(1+l)} \right)^\alpha \frac{R}{(\overline{s} - \underline{s})^\alpha} \sum_{i=0}^{m-1} \left(\frac{s^* - u}{\overline{s} - \underline{s}} \right)^i \frac{(i+\alpha)!}{i!}.$$

Theorem 4, due to Markoff, provides an uniform bound for the derivatives of a Lagrange polynomial.

Theorem 4 *(Markoff) Let $P_l(s) := \sum_j a_j s^j$ be a polynomial with real coefficients and degree l. If $\sup_{s \in S} |P_l(s)| \leq W$, then for all s in $intS$ and for all l in $\mathbb{N}$, it holds*

$$\left| P_l^{(j)}(s) \right| \leq \frac{l^2 (l^2 - 1) \ldots \left(l^2 - (j-1)^2 \right)}{(2j-1)!!} \left(\frac{2}{(\overline{s} - \underline{s})} \right)^j W.$$

When applied to the elementary Lagrange polynomial, it is readily checked that $W = \pi$. Indeed,

$$\begin{aligned} |L_{s_k}(s)| = & \left| \frac{(-1)^k \sin\left(\frac{2k-1}{2l+2}\pi\right)}{l+1} \frac{\cos((l+1)\theta)}{\cos\theta - \cos\left(\frac{2k-1}{2l+2}\pi\right)} \right| \leq \\ & \leq \frac{\left|\sin\left(\frac{2k-1}{2l+2}\pi\right)\right|}{l+1} \frac{|\cos((l+1)\theta)|}{\left|\cos\theta - \cos\left(\frac{2k-1}{2l+2}\pi\right)\right|} \leq \\ & \leq \frac{\left|\sin\left(\frac{2k-1}{2l+2}\pi\right)\right|}{l+1} \frac{(l+1)\left|\theta - \frac{2k-1}{2l+2}\pi\right|}{\frac{1}{\pi}\sin\left(\frac{2k-1}{2l+2}\pi\right)\left|\theta - \frac{2k-1}{2l+2}\pi\right|} = \pi. \end{aligned}$$

We used

$$|\cos((l+1)\theta)| = \left| \cos((l+1)\theta) - \cos\left((l+1)\frac{2k-1}{2l+2}\pi \right) \right| \leq (l+1)\left| \theta - \frac{2k-1}{2l+2}\pi \right|$$

and $\cos\left((l+1)\frac{2k-1}{2l+2}\pi\right)=0$. Moreover,

$$\left|\cos\theta-\cos\left(\frac{2k-1}{2l+2}\pi\right)\right|=2\sin\left(\frac{\theta+\frac{2k-1}{2l+2}\pi}{2}\right)\left|\sin\left(\frac{\theta-\frac{2k-1}{2l+2}\pi}{2}\right)\right|.$$

The concavity of the sine function on $[0,\pi]$ implies

$$\sin\left(\frac{\theta+\frac{2k-1}{2l+2}\pi}{2}\right)\geq\frac{1}{2}\left(\sin\theta+\sin\left(\frac{2k-1}{2l+2}\pi\right)\right)$$

$$\left|\sin\left(\frac{\theta-\frac{2k-1}{2l+2}\pi}{2}\right)\right|\geq\frac{2}{\pi}\left|\theta-\frac{2k-1}{2l+2}\pi\right|,\theta\in[0,\pi].$$

Remark 5 *The Cauchy theorem merely gives a rough upper bound. In order to obtain a sharper upper bound, we would assume some additional hypotheses on the derivatives of the function.*

3.2 Case 2: Y is an unbounded random variable

If the support of the random variable Y is not bounded and φ is a polynomial of unknown degree t, $t\leq g-1$, with g known, it's still possible to give an upper bound for the estimation error. Since

$$\varphi^{(d)}=\sum_{i=0}^{g-1}\frac{\varphi^{(d+i)}(s^*)}{i!}(u-s^*)^i=$$

$$=\sum_{i=0}^{g-1}\frac{\sum_{k=0}^{g-1}L_{s_k}^{(d+i)}(s^*)\varphi(s_k)}{i!}(u-s^*)^i=\sum_{k=0}^{g-1}L_{s_k}^{(d)}(u)\varphi(s_k),$$

$\varphi^{(d)}$ can be estimated as follows

$$\widehat{\varphi^{(d)}(u)}=\sum_{k=0}^{g-1}L_{s_k}^{(d)}(u)\overline{Y}(s_k).$$

We have in probability $\widehat{\varphi^{(d)}}\to\varphi^{(d)}$ for $\min(n_k)\to\infty$. So,

$$Var\left(\widehat{\varphi^{(d)}}\right)=\sum_{k=0}^{g-1}\left(L_{s_k}^{(d)}(u)\right)^2\frac{\varsigma}{n_k}\to 0,$$

where ς is the variance of Z. We use the Tchebycheff's inequality in order to obtain an upper bound for the estimation error. For a given η,

$$\Pr\left\{\left|\widehat{\varphi^{(d)}}-\varphi^{(d)}\right|\geq\eta\right\}\leq\frac{\sum_{k=0}^{g-1}\left(L_{s_k}^{(d)}(u)\right)^2\frac{\varsigma}{n_k}}{\eta^2}.$$

If we aim to obtain, for all fixed ω, $\Pr\left\{\left|\widehat{\varphi^{(d)}}-\varphi^{(d)}\right|\geq\eta\right\}\leq\omega$, we can choose n^{*} as the solution of the equation $\frac{\sum_{k=0}^{g-1}\left(L_{s_k}^{(d)}(u)\right)^2\frac{\varsigma}{n_k}}{\eta^2}=\omega$, that is

$$n^{*}=\frac{\sum_{k=0}^{g-1}\left(L_{s_k}^{(d)}(u)\right)^2\varsigma}{\omega\eta^2}.$$

The integer $[n^{*}]$ is such that the inequality $\Pr\left\{\left|\widehat{\varphi^{(d)}}-\varphi^{(d)}\right|\geq\eta\right\}\leq\omega$ is satisfied.
We remark that if we know the degree t of the polynomial, then it is sufficient to set $g-1=t$. When $\varphi(u)=\varphi^{d}(u)$ (i.e., $d=0$), we have $[n^{*}]=\frac{\sum_{k=0}^{g-1}\left(L_{s_k}(u)\right)^2\varsigma}{\omega\eta^2}$.
We underline that for $d=0$ and when t is known $\widehat{\varphi^{(d)}(u)}=\widehat{\varphi(u)}$ coincides with Hoel's estimator.
If the solely information on φ is that φ is analytic then we are constrained to give hypotheses on the derivatives of the function. More precisely, since $\operatorname{Im}\varphi\subseteq\mathbb{R}$, we can't apply the Cauchy theorem on the analytic functions; we can only say that $\varphi(v)=E(Y)\in\mathbb{R}$. So, we are not able to find a constant R such that $|\varphi(v)|\leq R$. Moreover, since we can't observe $\varphi(v)$ for $v\notin S$, we don't have any data to estimate M_{Taylor}.

4 Bibliography

References

[1] Bennett, G., 1962. Probability Inequalities for the Sum of Independent Random Variables, *Journal of the American Statistical Association*, 57, 297,33–45.

[2] Bernstein, S.N., 1918. Quelques remarques sur l'interpolation, *Math. Ann.*, 79, 1–12.

[3] Broniatowski, M., and Celant, G., 2007. Optimality and bias of some interpolation and extrapolation designs. *J. Statist. Plann. Inference*, 137, 858–868.

[4] Celant, G., 2003. Extrapolation and optimal designs for accelerated runs. *Ann. I.S.U.P.*, 47, 3, 51–84.

[5] Celant, G., 2002. Plans accélérés optimaux: estimation de la vie moyenne d'un système. *C. R. Math. Acad. Sci. Paris*, 335, 1, 69–72.

[6] Coatmélec, C., 1966. Approximation et interpolation des fonctions différentiables de plusieurs variables. *Ann. Sci. Ecole Norm. Sup.*, 83, 4, 271–341.

[7] Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58, 13–30.

[8] Hoel, P.G., Levine, A., 1964. Optimal spacing and weighting in polynomial prediction. *Ann. Math. Statist.*, 35, 1553–1560.

[9] Hoel, P.G., 1965. Optimum designs for polynomial extrapolation. *Ann. Math. Statist.*, 36, 1483–1493.

[10] Rivlin, T.J., 1969. An introduction to the approximation of functions. Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London

[11] Schönhage, A., 1961. Fehlerfortplanzung bei interpolation, *Numer. Math.*, 3, 62–71.

[12] Spruill, M.C., 1984. Optimal designs for minimax extrapolation. *J. Multivariate Anal.*, 15, 1, 52–62.

[13] Spruill, M.C., 1987. Optimal designs for interpolation. *J. Statist. Plann. Inference*, 16, 2, 219–229.

[14] Spruill, M.C., 1987. Optimal extrapolation of derivatives. *Metrika*, 34, 1, 45–60.

[15] Spruill, M.C., 1990. Optimal designs for multivariate interpolation. *J. Multivariate Anal.*, 34, 1, 141–155.